Paleoceanography and Paleoclimatology

RESEARCH ARTICLE
10.1029/2018PA003415

The Role of the Southern Ocean in Abrupt Transitions and Hysteresis in Glacial Ocean Circulation

Sophia K.V. Hines[1,2], **Andrew F. Thompson**[1], **and Jess F. Adkins**[1]

[1]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA, [2]Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

Key Points:
- Southern Ocean buoyancy forcing and vertical diffusivity distribution may give rise to hysteresis in the overturning circulation
- Hysteresis curves are sensitive to NADW density and Southern Ocean sea ice extent
- Circulation configuration ("figure-eight" versus "two-cell") is primarily sensitive to NADW density

Supporting Information:
- Supporting Information S1

Correspondence to:
S. K.V. Hines,
shines@ldeo.columbia.edu

Citation:
Hines, S. K., Thompson, A. F., & Adkins, J. F. (2019). The role of the Southern Ocean in abrupt transitions and hysteresis in glacial ocean circulation. *Paleoceanography and Paleoclimatology*, *34*, 490–510. https://doi.org/10.1029/2018PA003415

Received 6 JUN 2018
Accepted 4 MAR 2019
Accepted article online 15 MAR 2019
Published online 5 APR 2019

Abstract High-latitude Northern Hemisphere climate during the last glacial period was characterized by a series of abrupt climate changes, known as Dansgaard-Oeschger events, which were recorded in Greenland ice cores as shifts in the oxygen isotopic composition of the ice. These shifts in inferred Northern Hemisphere high-latitude temperature have been linked to changes in Atlantic meridional overturning strength. The response of ocean overturning circulation to forcing is nonlinear and a hierarchy of models have suggested that it may exist in multiple steady state configurations. Here, we use a time-dependent coarse-resolution isopycnal model with four density classes and two basins, linked by a Southern Ocean to explore overturning states and their stability to changes in external parameters. The model exhibits hysteresis in both the steady state stratification and overturning strength as a function of the magnitude of North Atlantic Deep Water formation. Hysteresis occurs as a result of two nonlinearities in the model—the surface buoyancy distribution in the Southern Ocean and the vertical diffusivity profile in the Atlantic and Indo-Pacific basins. We construct a metric to assess circulation configuration in the model, motivated by observations from the Last Glacial Maximum, which show a different circulation structure from the modern. We find that circulation configuration is primarily determined by North Atlantic Deep Water density. The model results are used to suggest how ocean conditions may have influenced the pattern of Dansgaard-Oeschger events across the last glacial cycle.

1. Introduction

Paleoceanographic records since the mid-1990s have shown evidence for multiple ocean circulation configurations across the glacial (Figure 1; Alley & Clark, 1999; Curry & Oppo, 2005; Oppo & Curry, 2012; Rahmstorf, 2002; Sarnthein et al., 1994). Differences between circulation at the Last Glacial Maximum (LGM) and in the modern ocean primarily involved rearrangement of water masses—specifically a shoaling of North Atlantic Deep Water (NADW) in the Atlantic during the LGM—without necessarily changing overturning circulation strength (Figures 1a vs. 1b in an Atlantic-only view of the ocean or Figures 1d vs. 1e in a global view of the ocean; Curry & Oppo, 2005). Circulation configuration changes earlier in the glacial, however, have been associated with changes in the strength of NADW formation (Figure 1c vs. other panels; Oppo & Curry, 2012).

During the middle of the last glacial period, there were a series of abrupt changes in Northern Hemisphere high-latitude climate, known as Dansgaard-Oeschger (DO) events, which were recorded in Greenland ice cores (Dansgaard et al., 1993; NGRIP members, 2004). These abrupt transitions in the climate system have been explained by a variety of mechanisms, including ice dynamics (Bassis et al., 2017; MacAyeal, 1993), ocean oscillations (Birchfield & Broecker, 1990; Bond et al., 1999; Broecker et al., 1990), and variability in the tropical atmosphere (Cane & Clement, 1999; Clement & Cane, 1999). It has also been suggested that threshold behavior in the climate system is more likely in intermediate climate states (Brook & Buizert, 2018; Buizert & Schmittner, 2015; Dome Fuji Ice Core Project Members: et al., 2017; McManus et al., 1999). Reconstructions of NADW formation strength appear to be in phase with these abrupt climate changes (Figure 2b; Henry et al., 2016; Böhm et al., 2015; Lippold et al., 2009; Lynch-Stieglitz, 2017; McManus et al., 2004).

It has long been recognized from various model representations of the ocean that the ocean overturning circulation has a nonlinear response to forcing and can exist in more than one steady state (Ferreira et al., 2011; Manabe & Stouffer, 1988; Rahmstorf, 2002; Stommel, 1961). One explanation for abrupt climate transitions is

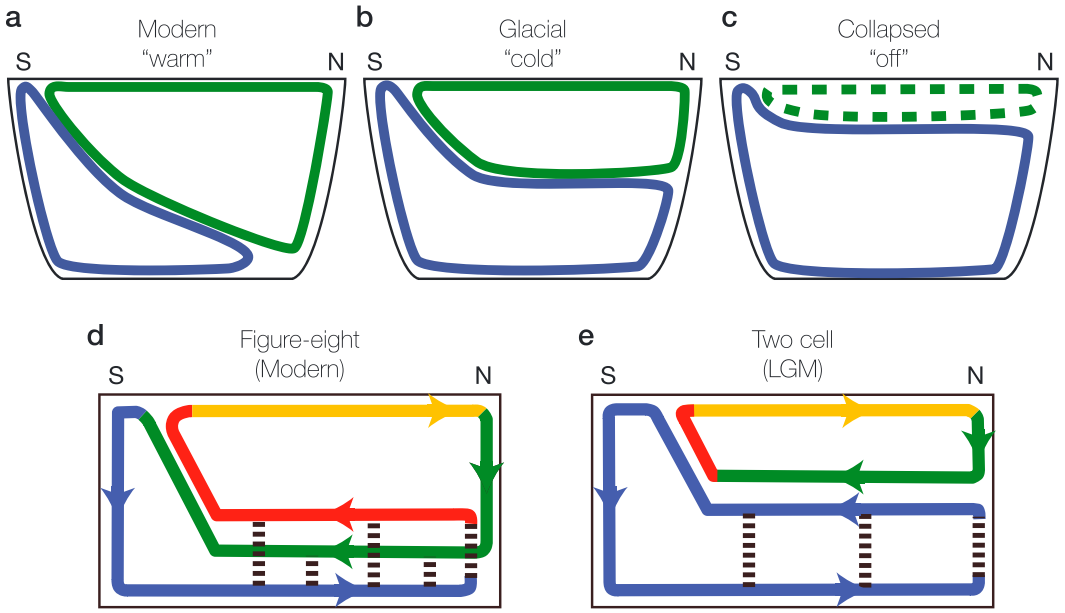

Figure 1. Schematic depicting various circulation modes postulated for the present and past ocean. (a–c) Depictions of Atlantic meridional overturning circulation across DO events (after Rahmstorf, 2002). In the modern "warm" state (a), NADW convects deeply in the North Atlantic, in the "cold" state (b), NADW does not penetrate as deeply in the basin, and in the "off" state (c), NADW collapses. (d, e) Depictions of the global meridional overturning circulation with the Atlantic and Pacific basins collapsed onto two dimensions (after Talley, 2013, and Ferrari et al., 2014). In the modern "figure-eight" configuration (d), circulation is closed diffusively in the Pacific basin, whereas in the "two-cell" configuration (e), circulation can close within a single basin. LGM = Last Glacial Maximum.

that they are the system response to ocean circulation transitioning from one steady state to another. Dynamical systems with multiple steady states may exhibit hysteresis, so these early models prompted a search for hysteresis in the overturning circulation. Heinrich Events, which precede the largest DO events, are associated with ice-rafted detritus (IRD) layers in marine sediments, formed by massive discharges of icebergs into the North Atlantic (Bond et al., 1993). These massive iceberg discharges result in large freshwater perturbations to the North Atlantic, which disrupt the formation of NADW. These events have been represented in coupled general circulation models (GCMs) and intermediate complexity earth system models by performing "hosing" experiments, where fresh water is added to a restricted domain in the North Atlantic for a short duration of time. Many of these experiments have exhibited hysteresis in the strength of Atlantic Meridional Overturning Circulation (AMOC) as a function of fresh water flux into the North Atlantic (Buizert & Schmittner, 2015; Rahmstorf et al., 2005). Other coupled ocean-atmosphere GCM models have demonstrated hysteresis in AMOC strength as a function of Northern Hemisphere ice volume and atmospheric CO_2 concentration (Zhang et al., 2014, 2017). Changes in overturning circulation in the Atlantic, however, must be linked with water mass transformations in other regions of the ocean in order to close the global overturning circulation.

The global overturning circulation involves a central role for the Southern Ocean (Kuhlbrodt et al., 2007; Marshall & Speer, 2012; Talley, 2013), where strong westerly winds give rise to the Antarctic Circumpolar Current (ACC). However, unlike the rest of the ocean, the ACC cannot support a mean geostrophic meridional flow because there are no continents to create sustained pressure gradients above the Drake Passage sill depth. Over the past decade, improved understanding of Southern Ocean dynamics and circulation has led to a theoretical framework that emphasized a near balance between the surface Ekman transport, that steepens isopycnals in response to a westerly wind stress, and eddies, that act to relax them. At steady state, the residual of these two components is balanced by water mass modification due to direct surface buoyancy forcing (Marshall & Radko, 2003). The wind-driven, or mean, component of the overturning circulation is defined as $\bar{\psi} = -\tau/\rho_0 f$, where τ is the surface wind stress, ρ_0 is the mean ocean density, and f is the Coriolis parameter. The counteracting eddy-driven component can be parameterized as $\psi^* = Ks_\rho$, where K is the eddy diffusivity (sometimes a function of slope) and s_ρ is the isopycnal slope. At steady state, the bal-

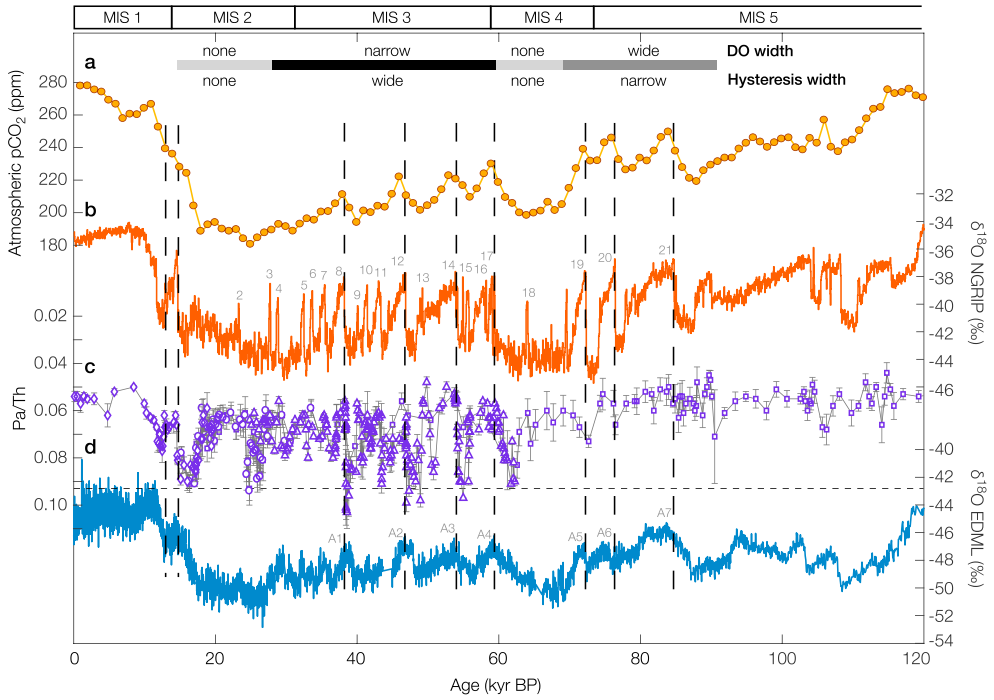

Figure 2. Atmospheric pCO_2, Greenland, and Antarctic $\delta^{18}O$ records with Pa/Th circulation tracer. (a) Atmospheric pCO_2 composite record from Antarctic ice cores (Bereiter et al., 2015). (b) Oxygen isotopic composition of NGRIP ice core in Greenland on the GICC05 time scale (Andersen et al., 2006; Rasmussen et al., 2006; Svensson et al., 2008; Vinther et al., 2006; Wolff et al., 2010) and (d) EDML ice core in Antarctica (EPICA Community Members, 2006). The $\delta^{18}O$ composition of ice is a proxy for regional temperature. Numbers mark the Dansgaard-Oeschger (DO) events in Greenland and corresponding AIM events in Antarctica. (c) Compilation of Pa/Th records from the Bermuda Rise (youngest section from McManus et al., 2004, diamonds; middle section from Lippold et al., 2009, circles; and Henry et al., 2016, triangles; oldest section from Böhm et al., 2015, squares). Pa/Th ratio is a proxy for the strength of North Atlantic Deep Water (NADW). Horizontal dashed line marks the Pa/Th production ratio of 0.093. Marine Isotope Stage boundaries are shown at the top, and gray bars denote the character of DO events during different time periods across the glacial.

ance between wind and eddy circulations leads to a structure where isopycnals slope from the surface to depth across the Southern Ocean, and the surface meridional distribution of density classes is mapped into a vertical stratification at the northern edge of the ACC (Nikurashin & Vallis, 2011). Importantly, though, this steady state isopycnal structure does not necessarily allow for net meridional transport—in fact if $\bar{\psi}$ perfectly matches ψ^*, then there is no net meridional overturning (for the adiabatic part of the residual circulation). The balance between the wind and eddy-driven circulation is known as the residual circulation: $\psi_{res} = \bar{\psi} + \psi^*$. The flow that is consistent with the residual overturning is the transport that is required to balance water mass transformation that occurs due to surface buoyancy forcing; transformation due to diapcynal mixing in the Southern Ocean are assumed to be small (Marshall & Radko, 2003; Speer et al., 2000). Where there is a positive (negative) surface buoyancy flux, water in the mixed layer can be converted from a heavier (lighter) density class to a lighter (heavier) density class. When the water mass transformation due to surface buoyancy forcing converges (diverges) mass into a given density class, it must be balanced by downwelling (upwelling) in this same density class. Because the residual circulation above the sill depth occurs along isopycnals in the ACC, it can permit an adiabtic pole-to-pole overturning along the subset of isopycnals that outcrop both in the Southern Ocean and in the North Atlantic (Wolfe & Cessi, 2011).

In the modern ocean, the closure of the overturning circulation requires water mass modification in the Indo-Pacific basins to balance the formation of dense waters in the North Atlantic and Southern Ocean (Lumpkin & Speer, 2007; Talley, 2013). This requirement is based on the observation that NADW upwells in a region of the Southern Ocean where it experiences a negative surface buoyancy forcing, and sinks again as Antarctic Bottom Water (AABW), part of the lower circulation cell. If this AABW flows back into the Atlantic basin, then it upwells back into NADW. It is only once AABW flows into the Indian and Pacific

basins that it is able to diffusively upwell to a density class where it can close the overturning loop. It does this either via Indian and Pacific Deep Water (IDW/PDW), which upwell in a positive buoyancy forcing region of the Southern Ocean and return north as Antarctic Intermediate Water (AAIW), or via the Indonesian throughflow, flowing first from the surface of the Western Equatorial Pacific into the Indian Ocean, and then from the Indian Ocean into the Atlantic through the Agulhas Leakage (Talley, 2013). In a zonally averaged depiction of the meridional overturning circulation, adiabatic pole-to-pole circulation is the upper circulation cell. The lower cell occupies density surface that only outcrop in the Southern Hemisphere and therefore must close through diabatic transformation, that is, diffusive mixing, in the ocean interior.

Building upon the work of Curry and Oppo (2005), Talley (2013), and Marshall and Radko (2003), Ferrari et al. (2014) recently put forth a theoretical model for changes in overturning circulation structure at the LGM compared to today. This model relies on two key observations: (i) that the change from positive surface buoyancy forcing to negative surface buoyancy forcing in the Southern Ocean aligns roughly with the quasi-permanent sea ice edge (defined as ice-covered 70% of the time), and (ii) that there is dramatically higher diapycnal mixing over rough bottom topography (Polzin et al., 1997). Ferrari et al. (2014) use a geometric argument to calculate the depth of the boundary between the upper and lower circulation cells in the northern basins based on the extent of sea ice in the Southern Ocean. For the modern, this depth comes out to ~2,200 m, below the mean depth of the mid-ocean ridges (2,000 m). If sea ice extent expanded at the LGM, as is implied by data and models, the boundary between the upper and lower circulation cells would shoal, raising it above the region of rough topography in the ocean, thus reducing diapycnal mixing. One potential result of this change in diapycnal diffusivity, driven by a shoaling of the boundary between the circulation cells, is a change from an intertwined "figure-eight" overturning circulation, as described by Talley (2013), into a two-cell overturning circulation (Curry & Oppo, 2005). Idealized model results from Jansen (2017) show a shoaling of NADW and an increase in separation between the upper and lower circulation cells in response to atmospheric cooling and subsequent Southern Ocean sea ice expansion, supporting aspects of this hypothesis.

In order to understand the complex interplay of physics and biogeochemistry that lead to glacial-interglacial cycles, we rely on models as well as paleoceanographic reconstructions, but it can be difficult to accurately model processes that occur on millennial time scales. On one end of the modeling continuum are GCMs, which attempt to capture all the physics of the ocean (and/or atmosphere) but are computationally expensive. It is widely understood that small-scale processes are crucial for accurately representing the Southern Ocean residual circulation and AABW formation, yet the computational cost is prohibitive to run high-resolution versions of these models, which explicitly resolve these processes for thousands of years. On the other end are simple theoretical models, which parameterize fluxes based on theoretical relations rather than attempting to resolve them based on the equations of motion. An early and well-known example of this type of model is from Gnanadesikan (1999), which was used to explain the depth of the pycnocline in the global ocean based on various water mass transformation processes. This model represented the ocean in a zonal average with two layers (two density classes). Subsequently, Shakespeare and Hogg (2012) included a lower overturning cell, focusing on responses to changing wind and buoyancy forcing in the Atlantic basin, and Marshall and Zanna (2014) expanded the Gnanadesikan model to multiple layers, but kept the same general model geometry and did not attempt to include AABW formation.

Our approach builds upon earlier models that have attempted to capture the key advances in residual circulation theory, most importantly that an imbalance between wind-driven and eddy overturning in the Southern Ocean is largely balanced by water mass transformation at the surface. Our study links this balance both to dynamics happening at low latitudes and our understanding of the surface buoyancy flux distribution in the Southern Ocean; this latter constraint has not previously been applied. Abernathey et al. (2016) have demonstrated that sea ice formation and sea ice melt dominate the surface buoyancy forcing in the Southern Ocean. The spatial separation between dense water formation (sea ice growth) over the shelf and light water formation (sea ice melt) further offshore demands that the overturning circulation is divided into upper and lower cells of different densities. In the lower cell, the loss of buoyancy at the surface in the Southern Ocean is balanced by buoyancy gain through vertical mixing in the basin interiors. In the upper cell, the gain of buoyancy at the surface in the Southern Ocean is balanced by a loss of buoyancy during NADW formation in the north. These essential features were recently codified in a model (Thompson et al., 2016) that allows the surface buoyancy distribution to evolve in response to a prescribed surface buoyancy flux. The model solves for the outcropping position of isopycnals at the surface in the ACC as well as their

depths in the basin at steady state, providing an explicit connection between basin stratification, isopycnal slope, and Southern Ocean surface processes. This model also keeps the Atlantic and Pacific basins separate, enabling it to occupy the figure-eight circulation configuration of Talley (2013). In this paper, we use this new time-dependent version of the Thompson et al. (2016) model to determine which conditions favor the various ocean configurations that have been observed and inferred for the past ocean (i.e., Figure 1). Although the simulations presented here are all run to steady state, the time dependence of the model allows us to find multiple equilibria, because we can approach these states from different directions. The simplicity of this model also permits an exploration of parameter space, including the effects of NADW strength, NADW density and the extent of sea ice in the Southern Ocean on the ability of the model to generate abrupt transitions and hysteresis in basin stratification. We investigate the origin of hysteresis, assess transitions in circulation configuration (figure-eight versus two-cell), and discuss the implications of our results for the interpretation of paleoceanographic records.

2. Model Description

We build upon the steady state model described in Thompson et al. (2016), keeping a similar architecture with four isopycnal layers and two basins, representing the Atlantic and Indo-Pacific (Figure 3). Each basin is two-dimensional and divided into two regions: a northern basin region, where interfaces between density classes are horizontal, and an ACC region, where isopycnals slope and outcrop at the surface. In the ACC region, mass may be transferred between the Atlantic and Indo-Pacific basins via zonal convergence and divergence (pink arrows in Figure 3). Note that this zonal convergence/divergence ($O(5)$ Sv) is more than an order of magnitude smaller than mean zonal flow ($O(150)$ Sv). The model is forced by an imposed buoyancy flux and wind stress at the surface of the Southern Ocean (plotted as "F_b" and "τ," respectively, in the upper left of Figure 3), and imposed NADW forcing in the northern end of the Atlantic basin (green arrow in Figure 3). The density of NADW, ϕ, determines the fraction of NADW that goes into layer 2 versus 3 (split in green arrow). Based on these imposed forcings, the model adjusts to find a steady state circulation and stratification. From initial conditions, the model evolves the depth of each isopycnal in the basin, z, and each outcrop position in the ACC, y, until a steady state configuration is achieved. The evolution equations are given by

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{1}{A}\left[\Delta T_{\mathrm{diff.}} + \Delta T_{\mathrm{res.}} + \Delta T_{\mathrm{zonal}} + \left(T_{\mathrm{NADW}}\right)\right], \tag{1}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{h_m L_{A,P}}\left(\bar{\psi} + \psi^* - \psi_{\mathrm{res.}}\right) = \frac{1}{h_m}\left(-\frac{\tau}{\rho_0 f} + K s_\rho - \left.\frac{F_b}{\partial_y b}\right|_{\mathrm{surf.}}\right). \tag{2}$$

Equation (1) is a volume budget for the part of the ocean bounded by z and the interface above: $V = (z_i - z_{i-1}) * A$, where A is the horizontal area of the basin (taken as a constant). In Equation (1), the ΔT terms represent net transports: the net diffusive transport, $\Delta T_{\mathrm{diff.}}$, is dependent on a vertical diffusivity ($\kappa(z)$) times the stratification ($\partial_z \rho$), the net residual transport, $\Delta T_{\mathrm{res.}}$, is dependent on the surface buoyancy flux (F_b) and the meridional surface buoyancy gradient ($\partial_y b$), and the net zonal transport, $\Delta T_{\mathrm{zonal}}$, is dependent on the difference in cross-sectional area of a given layer in the depth latitude plane between the Atlantic and Pacific basins. T_{NADW} is the imposed NADW flux given in Sverdrups (1 Sv = 10^6 m^3/s). Stratification is directly related to layer thickness: high stratification corresponds with thin layers and low stratification corresponds with thick layers. More detailed equations are provided in the Appendix.

The outcrop position of an isopycnal, y, is determined according to residual mean theory as parameterized in equation (2). At steady state ($\mathrm{d}y/\mathrm{d}t = 0$), it simplifies to the relation $\psi_{\mathrm{res}} = \bar{\psi} + \psi^*$. The terms h_m and $L_{A,P}$ represent the depth of the mixed layer and width of the Atlantic or Pacific basin, respectively. The parameter τ is the wind stress, ρ_0 is the reference density, f is the Coriolis parameter (set to a constant value of -1×10^{-4} s^{-1}, which is representative of the Southern Ocean), K is the along-isopycnal diffusivity, s_ρ is the isopycnal slope ($-z/y$; always negative), F_b is the surface buoyancy flux, and $\partial_y b$ is the surface buoyancy gradient. The depths and outcrop positions for each isopycnal can differ in the Atlantic and Pacific. See Appendix A for a detailed description of model equations, definitions of the various transformation processes, and parameterizations.

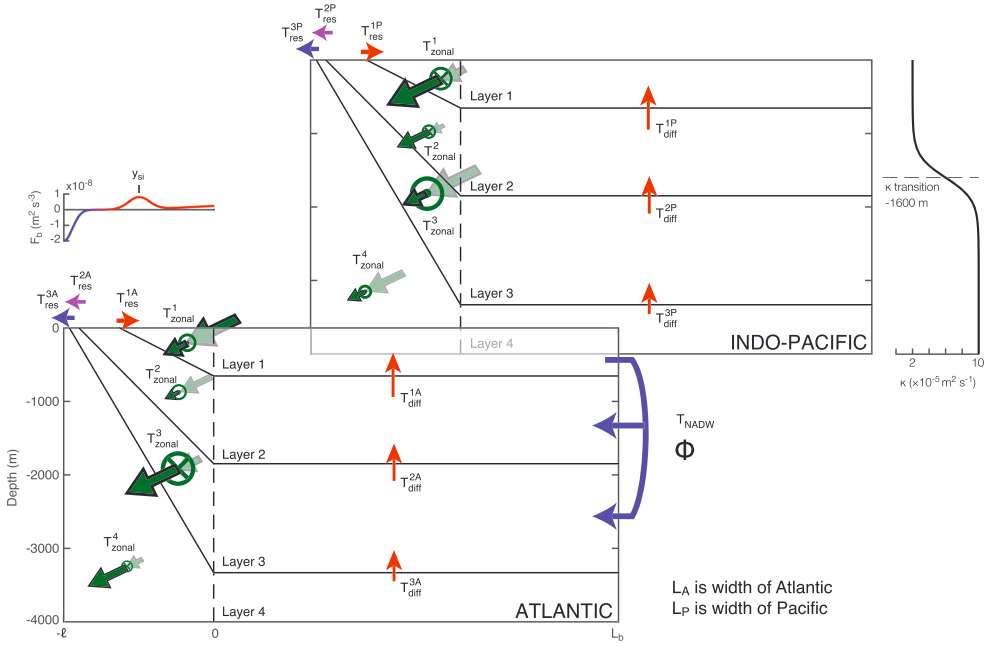

Figure 3. Schematic of model architecture. Model domain is split into an Atlantic sector and an Indo-Pacific sector. Each basin has a channel region with sloping isopycnals that outcrop at the surface (from $-\ell$ to 0), and a basin region where isopycnals are flat lying (from 0 to L_b). Volume fluxes are marked by arrows that are colored based on the direction of the transformation: blue indicates that light water is made heavier, red indicates that heavy water is made lighter, and green indicates that water does not change density. Purple arrows (for $T_{\text{res}}^{2A,P}$) signify that this flux can be either light to heavy or heavy to light depending on the simulation. Within each basin, water can move between density classes due to diabatic processes at the surface of the channel region ($T_{\text{res}}^{iA,P}$) or due to diapycnal diffusion in the basin region ($T_{\text{diff}}^{iA,P}$). The profile of vertical diffusivity, κ, is shown to the right of the Pacific basin (see equation (A4)). Water moves between the Atlantic and Indo-Pacific due to zonal convergence in the channel (green arrows). For zonal exchange, $\otimes$ represents divergence out of the basin and $\odot$ represents convergence into the basin. Note that in the ACC zonal convergence/divergence ($O(5)$ Sv) is much smaller than mean zonal flow ($O(150)$ Sv). At the northern end of the Atlantic basin, North Atlantic Deep Water is formed with variable density (blue arrow; density specified by parameter ϕ). The plot in upper left shows surface buoyancy flux, F_b, as a function of y for the channel region of the model with sea ice edge, y_{si}, marked (red color indicates positive buoyancy flux and blue color indicates negative buoyancy flux).

Each layer in the model has a fixed buoyancy ($b \equiv -g(1-\rho/\rho_0)$), which means that as layer volumes adjust, the buoyancy content of the ocean changes as well. Buoyancy fluxes to the ocean are explicitly represented at the surface of the Southern Ocean, but they are implicitly involved in our imposed NADW flux forcing and diffusive volume fluxes in the basin. Once a steady state is achieved in the model, the implicit buoyancy fluxes involved in these Northern Hemisphere and low-latitude water mass transformations are balanced by the net buoyancy flux in the Southern Ocean, so the buoyancy content of the ocean is constant and in balance. This surface buoyancy flux in the North Atlantic is not explicitly provided in our model because the relationship between this flux and water mass transformation (which is a volume flux between density classes) is not well understood in this region. Therefore, we choose to prescribe the volume flux and to focus on the volume budget rather than the buoyancy budget.

The distribution of surface buoyancy forcing in our model is idealized, but motivated by a critical aspect of the climate system—Abernathey et al. (2016) and Pellichero et al. (2018) have shown that sea ice distributions and associated buoyancy fluxes dominate the total surface buoyancy flux and water mass transformations in the Southern Ocean. As long as sea ice is present, there will be a negative buoyancy flux region near the Antarctic continent associated with brine rejection and net sea ice formation, a positive buoyancy flux region near the sea ice edge associated with fresh water from net sea melt, and a region covered by sea ice with very little surface buoyancy flux. The surface buoyancy flux in our model matches this pattern, with a small additional positive buoyancy flux north of the sea ice edge representing atmospheric heating.

Likewise, our vertical diffusivity profile is also based on key direct observations, which show that vertical diffusivity is enhanced over regions of rough bottom topography in the ocean (Polzin et al., 1997; Waterhouse

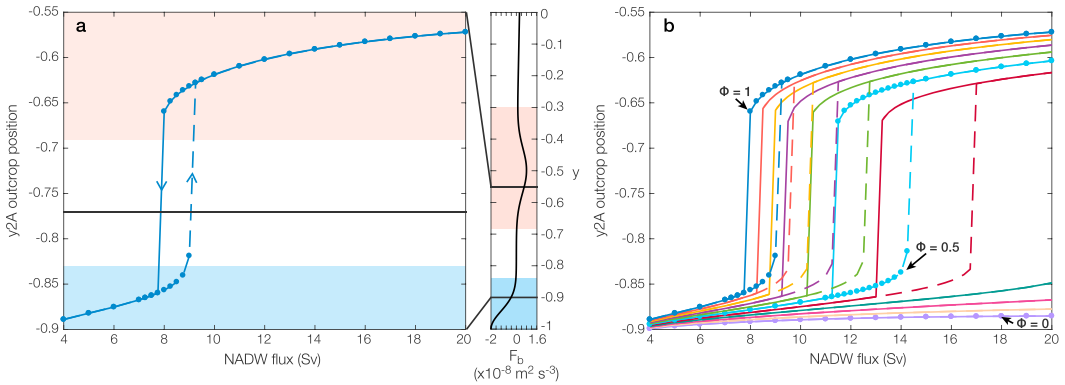

Figure 4. Outcrop position of interface 2 in the Atlantic (y_2^A) for all values of ϕ (NADW density). Panel (a) shows a single hysteresis loop for a ϕ value of 1 (all NADW goes into layer 3). Shading in the background indicates regions of strong positive and negative surface buoyancy flux, F_b (plotted in full to right). Horizontal black line marks the $F_b = 0$ contour. Panel (b) shows all hysteresis loops for ϕ varying from 0 to 1 in increments of 0.1. Each loop is a different single color; dashed lines mark the upgoing limb of the loop (i.e., increasing from 4 to 20 Sv) and solid lines mark the downgoing limb of the loop.

et al., 2014). At a given location, vertical diffusivity tends to increase exponentially toward the bottom, but there is a distribution of rough topography in depth, with the tops of the mid-ocean ridges occurring globally at roughly 2,000-m water depth and almost 90% of the seafloor area occurring below this depth (de Lavergne et al., 2017). As a result, there will always be enhanced diapycnal mixing and diffusive transport in the deep ocean compared to the upper ocean. This distribution is represented in our model by strong vertical diffusivities at depth and weak diffusivities in the upper ocean with a transition region that is centered at 1,600 m for most experiments, but can be varied (equation (A4) in the appendix).

3. Results

We have explored various steady state configurations of the overturning circulation over a wide range of parameter space. In our experiments, we varied the strength of NADW formation, T_{NADW}, the density of NADW, ϕ, and the position of the sea ice edge, y_{si}. We also examined how the vertical profile of the vertical diapycnal diffusivity in the basin impacts the steady states for the overturning circulation. Specifically, we wanted to assess which conditions supported abrupt changes in basin stratification due to perturbations in the strength of T_{NADW}. Furthermore, we diagnosed which conditions were associated with a "figure-eight" versus "two-cell" circulation configuration.

Abrupt rearrangements in the overturning circulation involve changes in the stratification and overturning strength throughout the model. These whole-ocean changes lie between two endmember states. One endmember is associated with low NADW flux, with enhanced stratification (thinner layers) in the deep ocean and isopycnals that outcrop further to the south in the Southern Ocean. This in turn enhances deep circulation, since more of the isopycnals outcrop in the negative buoyancy forcing region of the Southern Ocean. In the other endmember state, associated with high NADW flux, there is weaker stratification (thicker layers) in the deep ocean and isopycnals that outcrop further to the north in the Southern Ocean. In this configuration, more isopycnals outcrop in the positive buoyancy forcing region of the Southern Ocean and the upper adiabatic cell is strengthened.

Here we focus our discussion on the Southern Ocean outcrop position of the interface between layers 2 and 3 in the Atlantic basin, y_2^A. This interface is critical to shifts in the overturning structure because it determines how NADW, upwelling in the Southern Ocean, is partitioned between AABW and AAIW. In addition to exploring the sensitivity of the position y_2^A to NADW formation strength, we also considered the impact of NADW density, parameterized by ϕ, where ϕ takes values between 0 (NADW formation involves a transformation from layers 1 to 2) and 1 (NADW formation involves a transformation from layers 1 to 3). The value of ϕ is changed in increments of 0.1, where ϕ indicates the fraction of NADW entering layer 3. For each value of ϕ a suite of experiments was carried out in which the overturning was allowed to evolve to a steady state for a given value of T_{NADW}. This steady state configuration was then used as an initial condition for the next simulation where T_{NADW} was modified; T_{NADW} was increased monotonically from 4 to 20 Sv,

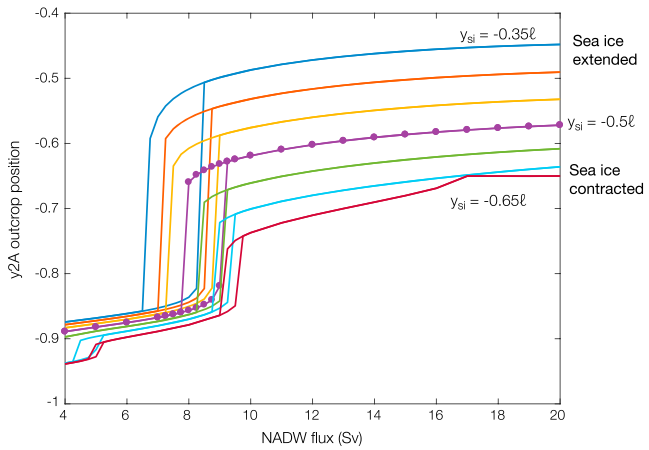

Figure 5. Outcrop position of y_2^A as a function of NADW flux for different sea ice extents (y_{si}). A y_{si} value of −0.5 means that the maximum of F_b associated with sea ice melt is centered at the middle of the ACC and is an equivalent position to −0.5 on the y axis (see equation (A8)). From bottom to top, y_{si} position increases in increments of 0.05. NADW density (ϕ) was held constant for all experiments at a value of 1.

and then back to 4 Sv; the T_{NADW} step size was reduced close to where abrupt transition in the overturning structure was found (Figure 4).

For values of ϕ between 0.4 and 1, the suite of steady state overturning configurations, the NADW loop, is characterized by an abrupt transition. In particular, this results in the an abrupt shift in y_2^A from a region of negative surface buoyancy forcing to a region of positive surface buoyancy forcing (Figure 4a). These steady state solutions also exhibit hysteresis in y_2^A as a function of T_{NADW}—the transition from outcropping in a negative buoyancy flux region to a positive buoyancy flux region occurs for larger values of T_{NADW} when increasing T_{NADW} than when T_{NADW} is being reduced (Figure 4). The width of the hysteresis, given in Sv increases as ϕ decreases from 1 to 0.4, but once ϕ is below 0.4, the hysteresis (and the abrupt transition in outcrop position) disappears (Figure 6c).

To test the influence of sea ice extent on the circulation configuration, we varied y_{si} between -0.35ℓ (expanded sea ice extent) and -0.65ℓ (contracted sea ice extent) with ϕ held constant at a value of 1 (see Figure 3). Modifying y_{si}, while changing the meridional distribution of the surface buoyancy flux F_b, does not change the total, or integrated, flux of buoyancy into the Southern Ocean. Specifically, changes to y_{si} modifies the meridional extent of the region where F_b is small (see equation (A8) in Appendix A). As sea ice extends northward (southward), the magnitude of the abrupt transition in y_2^A increases (decreases; Figure 5). The hysteresis in y_2^A outcrop position occurs for all values of y_{si} considering in our simulations (Figure 5).

For the two simulation suites where y_{si} is located furthest to the south ($y_{si} = -0.65\ell$ and -0.6ℓ), a second, smaller hysteresis loop occurs at low T_{NADW} values. This loop results from y_1^A transitioning between positive and negative buoyancy flux regions in the Southern Ocean. In this study, we focus only on the hysteresis loop that forms as the result of y_2^A transitioning between positive and negative buoyancy flux regions. However, it is interesting to note that y_1^A outcropping in the negative buoyancy flux region of the Southern Ocean implies that the only source of value transport back to the uppermost layer (layer 1) occurs through diffusion in the basin. This would be a purely diffusive overturning with no role for an "adiabatic" upper cell, as seen in the modern day, that closes primarily due to buoyancy gain in the Southern Ocean.

The width of the hysteresis loop, defined as the difference in NADW volume flux between the experiments bounding the region where multiple steady states occur, is dependent both on the density of NADW (ϕ) and the meridional extent of sea ice in the Southern Ocean (y_{si}). The maximum hysteresis width occurs at intermediate NADW densities ($\phi = 0.4$) and maximum northward sea ice extent ($y_{si} = -0.36\ell$; Figure 6c). Changes in hysteresis width are more sensitive to ϕ than y_{si}. Widths range from 0.75 Sv for $\phi = 1$ and $y_{si} = -0.65\ell$ to 5.5 Sv at $\phi = 0.4$ and a $y_{si} = -0.35\ell$. The NADW value at the center of the hysteresis loop also changes as a function of y_{si} and ϕ (Figure 6d). The T_{NADW} value at the center of the hysteresis loop is also more sensitive to ϕ than to y_{si}. Hysteresis center values range from 7.5 Sv at $\phi = 1$ and $y_{si} = -0.35\ell$ to 16.4 Sv at $\phi = 0.4$ and $y_{si} = -0.65\ell$.

4. Discussion

The primary goal of these experiments was to elucidate the physics giving rise to abrupt transitions in circulation during the last glacial period, with a particular emphasis on the link between surface buoyancy distribution and water mass transformation in the Southern Ocean. Although the model is fairly simple, it includes many key aspects of ocean physics, including both Atlantic and Indo-Pacific basins, a representation of the residual circulation in the Southern Ocean, a depth-dependent vertical diffusivity in the northern basins, and a Southern Ocean buoyancy forcing constrained by freshwater fluxes from Antarctic sea ice formation and melt.

4.1. The Origin of Hysteresis and Abrupt Transitions in the Model

One of the most notable and interesting results of our experiments is the emergence of hysteresis in the overturning circulation. Similar to other models that have indicated abrupt transitions in the overturning

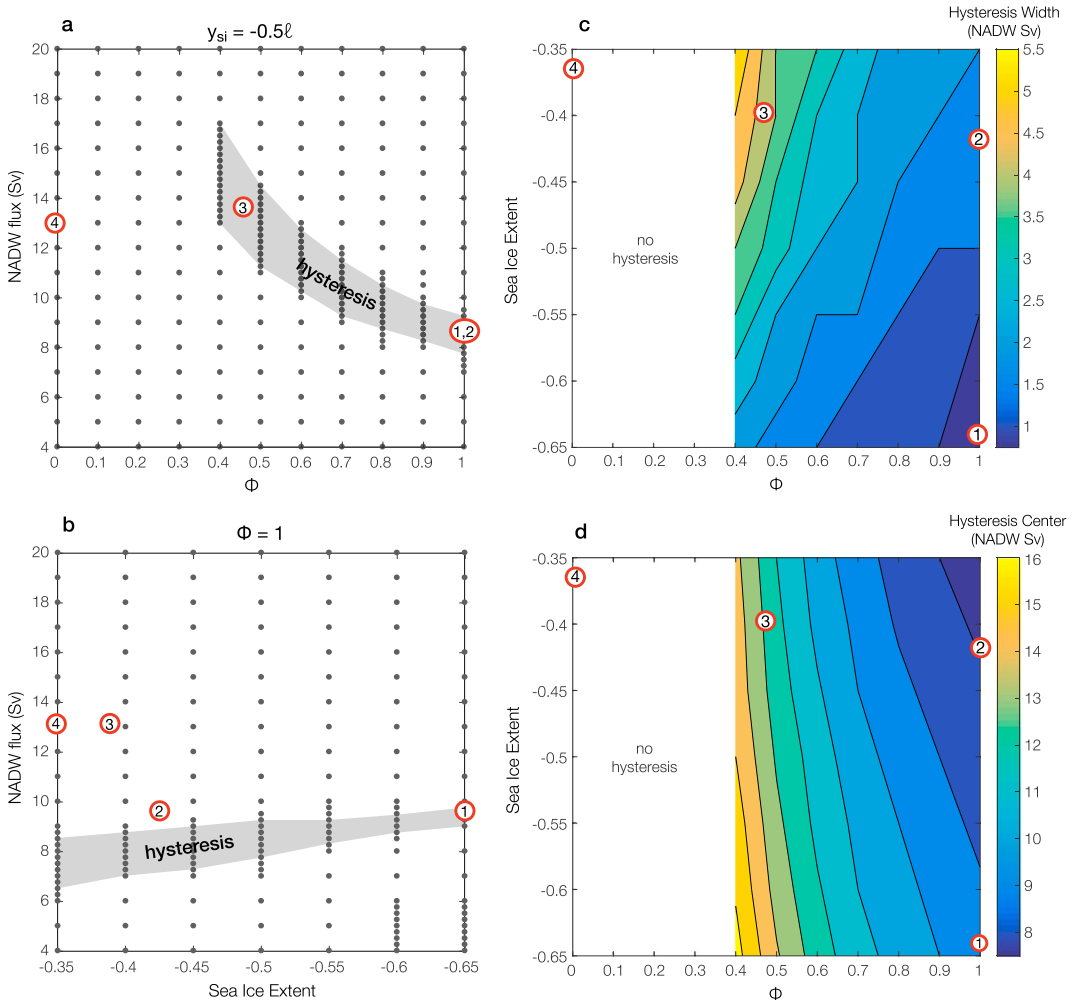

Figure 6. Hysteresis width and center as a function of y_{si} and ϕ. (a) NADW flux plotted versus ϕ for a y_{si} value of -0.5ℓ. Each dot denotes an experiment that was run, and the gray field indicates where there are multiple steady states. The "width" of the hysteresis is the vertical length of the gray bar for a particular value of ϕ, and the "center" of the hysteresis is the NADW value at the middle of the gray bar for a particular value of ϕ. (b) NADW flux plotted versus y_{si} for a ϕ value of 1. Note that panels (a) and (b) are projections of a three-dimensional space (NADW flux versus ϕ versus y_{si}) onto two dimensions. (c) The width of the hysteresis loop (in Sv of NADW) is contoured as a function of NADW density, ϕ, and sea ice extent, y_{si}. (d) The center point (in Sv of NADW) of the hysteresis loop is contoured as a function of NADW density, ϕ, and sea ice extent, y_{si}. A ϕ value of 1 indicates that all NADW flux goes into layer 3 and a ϕ value of 0 indicates that all NADW flux goes into layer 2. A y_{si} value of -0.5 means that the maximum of F_b associated with sea ice melt is centered at the middle of the ACC, and sea ice extends northward as the value of y_{si} increases. Areas of the contour plot without shading indicate that there is no hysteresis. Numbers in red circles refer to different times during the glacial (see section 4.6).

structure to small changes in NADW formation rates, here the focus is not only on the AMOC but rather the whole ocean adjustment. We are particularly interested in how characteristics of Southern Ocean sea ice and its associated surface buoyancy flux set the characteristics of the hysteresis loop. In dynamical systems, hysteresis indicates the presence of multiple steady states, and rapid flips between these states can give rise to abrupt transitions in the system. During Marine Isotope Stage (MIS) 3, there were many such abrupt transitions in the climate system, as recorded in isotope records from the Greenland ice sheet (Figure 2b).

This hysteresis in ocean density structure is captured by monitoring the Southern Ocean outcrop position of y_2^A, which abruptly jumps between regions of positive and negative buoyancy forcing. At steady state, the surface buoyancy forcing in the Southern Ocean imposes a constraint on the slope of each interface ($s \equiv -z/y$; always negative), given its outcrop position. Between regions of ice formation and melt, the surface buoyancy forcing is weak ($F_b \approx 0$). For a density surface to outcrop in this region, a steady state configuration

Table 1
Values and Definitions of Parameters Used in Model

Parameter	Value	Description
L_y	1×10^7 m	Meridional basin length
ℓ	2×10^6 m	Meridional ACC length
L_A	3×10^6 m	Width of Atlantic basin
L_P	10×10^6 m	Width of Pacific basin
H	4,000 m	Ocean depth
h_m	100 m	Depth of Southern Ocean mixed layer
τ_0	0.1 N/m^2	Reference wind stress
f	-1×10^{-4} s^{-1}	Coriolis parameter
ρ_0	1,025 kg/m^3	Reference seawater density
F_0	2×10^-8 m^2/s^3	Reference surface buoyancy flux
τ_r	$2 * 2.6 \times 10^6$ s	Relaxation time scale for atmospheric heat flux (2.6×10^6 s is equal to 1 month)
Δb_{atm}	1.3×10^{-3} m/s^2	Range of atmospheric buoyancy forcing (choice of value corresponds to atmospheric heat flux of 50 W/m^2 at northern boundary of ACC.)
b_1, b_2, b_3, b_4	$(5, 1, -1, -2.5) \times 10^{-3}$ m/s^2	Buoyancy of AAIW, UCDW/PDW, LCDW/NADW, and AABW density classes, respectively
a_N, a_S	$(2, 1.6) \times 10^5$ m	Widths of the sea ice melting and formation regions, respectively
K	1000 m^2/s	ACC along-isopycnal diffusivity
κ_0	1×10^{-4} m^2/s	Mean diapycnal diffusivity
$\Delta\kappa$	8×10^{-5} m^2/s	Change in diapycnal diffusivity
z_0	1,600 m	Depth of transition in diapycnal diffusivity
d	300 m	Thickness of transition in diapycnal diffusivity
U_0	0.05 m/s	Mean (barotropic) zonal velocity
r_b	1	Parameter for relative bottom water formation in Atlantic versus Pacific
T_{NADW}	$4–20 \times 10^6$ m^3/s	Forcing parameter: NADW flux (varied in experiments)
ϕ	0–1	Forcing parameter: NADW splitting parameter (varied in experiments)
y_{si}	$-(1.3–0.7) \times 10^6$ m	Forcing parameter: Meridional position of sea ice edge in Southern Ocean (varied in experiments)

Note. ACC = Antarctic Circumpolar Current; AAIW = Antarctic Intermediate Water; NADW = North Atlantic Deep Water; PDW = Pacific Deep Water; AABW = Antarctic Bottom Water; UCDW = Upper Circumpolar Deep Water; LCDW = Lower Circumpolar Deep Water.

implies that wind and eddy contributions to the overturning must exactly balance. In the absence of a surface buoyancy flux, no water mass transformation can occur, and thus there is no volume transport between density classes. For this case, we can define a critical slope, s_0, using the evolution equation for y at steady state (equation (2)):

$$\left.\frac{F_b}{\partial_y b}\right|_{\text{surf.}} = -\frac{\tau}{\rho_0 f} + K s_\rho. \tag{3}$$

$$\text{For} \quad F_b = 0, \quad s_0 \equiv \frac{\tau}{\rho_0 f K} = -9.8 \times 10^{-4}, \tag{4}$$

values for parameters in equations (3) and 4 are given in Table 1. At steady state, there is a strong constraint on the slope of an isopycnal from equation (3). If the isopycnal outcrops in a positive buoyancy forcing region, then s_ρ must be shallower (or less negative) than s_0, if the isopycnal outcrops in a negative buoyancy forcing region, then s must be steeper (or more negative) than s_0. Note that s is always negative. For the experiment suite with $\phi = 1$ (Figure 4), the hysteresis loop is centered on the critical slope $s_\rho = s_0$ for y_2^A (Figure 7).

Within a traditional hysteresis loop, there are three steady state solutions, two of which are stable (the stable solutions here are determined numerically). From Figure 7, we see that the two stable solutions are where

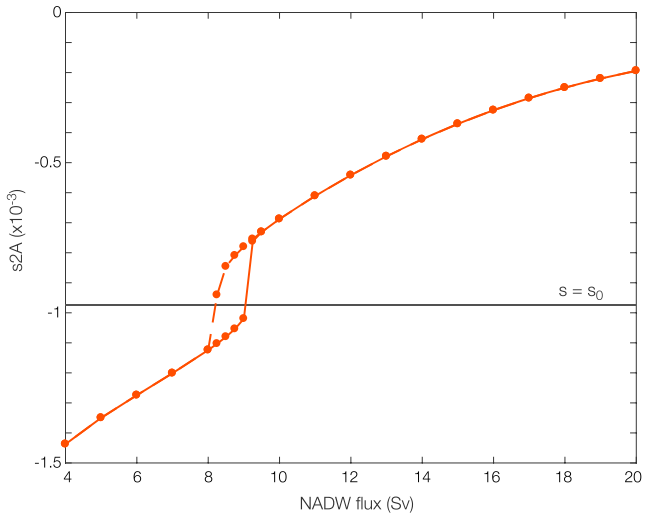

Figure 7. Interface 2A slope as a function of North Atlantic Deep Water (NADW) flux for $\phi = 1$. The slope for interface 2 in the Atlantic is defined as $-z_2^A/y_2^A$. Horizontal line marks $s = s_0$.

$s > s_0$ and $s < s_0$. A third solution would typically lie between the two stable solutions, presumably where s is equal to s_0 and F_b equals zero. This solution is unstable because any perturbation to the isopycnal outcrop position moves it to a region with a larger buoyancy flux (either positive or negative), which then requires the slope to adjust further. This interpretation should be taken with some caution, though, as a steady state configuration requires all evolutions equations (for both outcrop positions and interface depths in the basin) to be steady. The tendency to move away from a regime where $s_\rho = s_0$ may be triggered by a change in the low-latitude stratification.

While the surface buoyancy flux distribution in the Southern Ocean plays a large part in the emergence of the hysteresis loop, it is not solely responsible for generating hysteresis. The nonlinear distribution of vertical diffusivity in the basin, $\kappa(z)$, also contributes (see Figure 3). Indeed, running the same NADW loop with a constant or linear κ profile results in the disappearance of hysteresis and abrupt transitions in outcrop position (Figure 8).

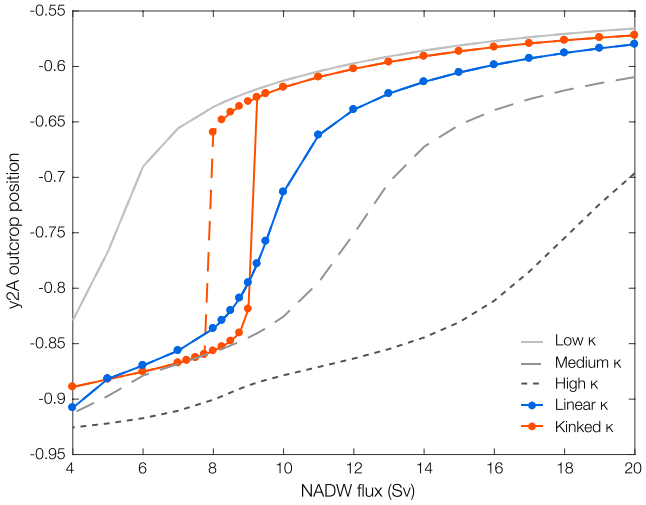

Figure 8. Outcrop position of y_2^A as a function of North Atlantic Deep Water (NADW) flux for different vertical diffusivity profiles. The control experiment has a kinked κ profile (red). A series of experiments with constant κ (at values of 0.2×10^{-4}, 0.6×10^{-4}, and 1×10^{-4}; gray) and an experiment with a linearly increasing κ profile (from 0.2×10^{-4} to 1×10^{-4}; blue) are shown for comparison.

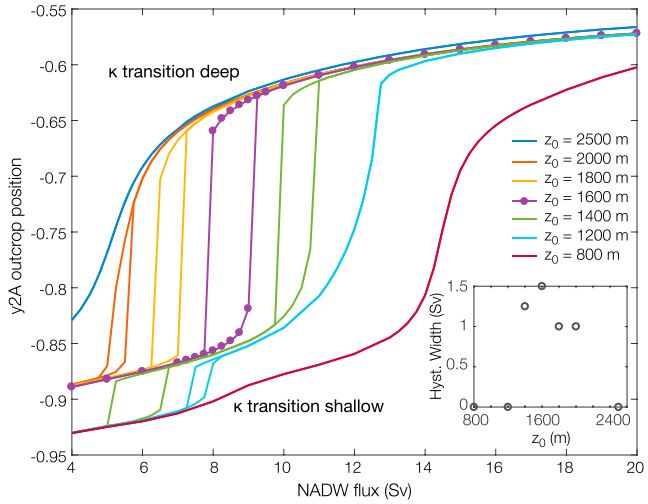

Figure 9. Outcrop position of y_2^A as a function of North Atlantic Deep Water (NADW) flux ($\phi = 1$) for different κ transition depths, z_0. The width of the hysteresis in the outcrop position of y_2^A as this interface flips between the positive and negative buoyancy flux regions is dependent on the κ transition depth and has a maximum at $z_0 = 1{,}600\text{m}\%$ (see inset plot).

In fact, there is an interplay between the nonlinear surface buoyancy distribution and the nonlinear vertical diffusivity. If the "kink" in the diffusivity profile, the depths over which κ transitions from strong to weak values, is too shallow or too deep, the steady state solutions are similar to those obtained with a uniform value of κ and the hysteresis loop disappears (Figure 9). The inset plot of hysteresis width as a function of kink depth shows that a depth of 1,600 m produces the maximum hysteresis width.

To understand how the κ profile contributes to the emergence of hysteresis, we consider a simplified system with a single basin and only two density classes (two layers), separated by a single interface. The diffusive transport, T_{diff}, is strongly modified as the interface moves through the κ transition region. In this simplified system, we can analytically solve for the diffusive transport as a function of interface depth for a variety of κ profiles (Figure 10). The diffusive transport is defined as $T_{\text{diff.}} = \kappa(z)A/\Delta z$, where Δz is the upper layer thickness, equal to the depth of z in the simplified two-layer model (Figure 10a). Figure 10c shows that the nonlinear, "kinked" κ distribution produces the same diffusive transport at multiple depths; uniform or linear profiles lead to monotonic changes in T_{diff}. Multiple depths z for the interface that lead to similar T_{diff} values allows the ocean interior to match Southern Ocean surface transformations under very different stratifications that may help to accommodate large changes in the surface buoyancy distribution in the Southern Ocean as y_2^A reaches a new, stable steady state position. Note that the low diffusive transport values

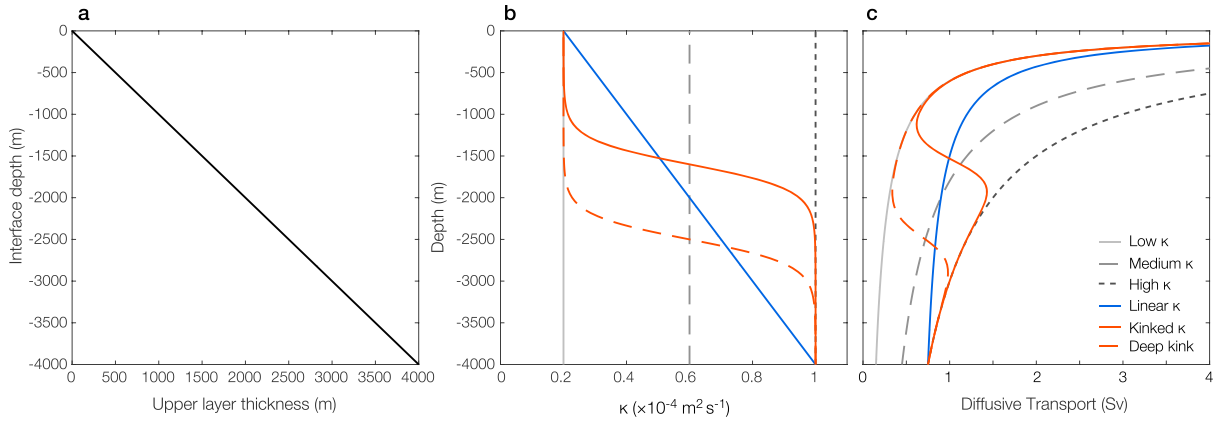

Figure 10. Calculated diffusive transport for a variety of κ profiles. (a) Thickness of layer 1 as a function of interface depth. (b) Vertical diffusivity, κ, as a function of depth. (c) Diffusive transport, defined as $T_{\text{diff}} = (\kappa A) / (\Delta z)$, where Δz is the layer thickness and A is the horizontal area of the basin, for the various κ profiles. "Deep kink" has the κ transition at 2,500 m.

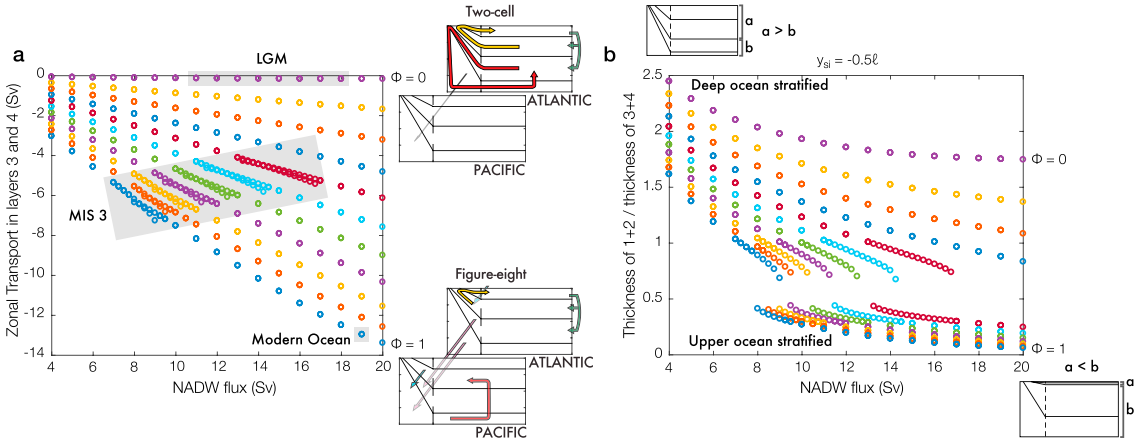

Figure 11. Assessment of circulation configuration and stratification for all ϕ experiments. (a) Combinations of NADW strength and density that have very little zonal transport in layers 3 and 4 lead to a circulation configuration that is two-cell-like, and combinations of North Atlantic Deep Water (NADW) strength and density that have large zonal transport in layers 3 and 4 lead to a circulation configuration that is more figure-eight-like. Different colored circles represent different ϕ values from 0 (at top) to 1 (diagonal to lower right corner). Shaded regions show location of the modern ocean and approximate locations for the Last Glacial Maximum (LGM) and Marine Isotope Stage (MIS) 3. (b) Ocean stratification, defined as the ratio of the average thickness of upper ocean layers divided by the average thickness of deep ocean layers. Hysteresis shows up strongly in this parameter.

(around 1 Sv) are a result of reducing the model to a single basin and two layers—T_{diff} values in the four-layer model are ~1–8 Sv for a NADW flux of 15 Sv.

4.2. What Defines the Limits of Hysteresis?

The interplay between surface nonlinear buoyancy forcing and nonlinear vertical diffusivity, which gives rise to hysteresis in our model, also plays a role in determining the limits of where we find hysteresis. In our experiments, hysteresis disappears when ϕ drops below 0.4 (when less than 40% of NADW flux goes into layer 3 and more than 60% of NADW flux goes into layer 2). As a larger fraction of NADW flux enters layer 2 rather than layer 3 the depth of the interface that separates these two layers becomes less sensitive to the NADW flux (see equation (A11) and supporting information Figure S3). While this reduced sensitivity makes a transition less likely, it does not explain the limit of $\phi = 0.4$ for the presence of hysteresis.

Based on two features of the system, (i) the increase in depth of z_2^A as ϕ decreases (this occurs because more mass is injected into layer 2) and (ii) the emergence of hysteresis when the key interface adjusts across the κ transition zone, we tested the sensitivity of the hysteresis loop to the vertical position of the kink in κ. Indeed, deepening the kink from 1,600 to 1,800 m increases the range of ϕ values over which hysteresis occurs; hysteresis loops are now found for $0.3 < \phi < 1.0$, rather than for $0.4 < \phi < 1.0$ (see Figure S4).

With regard to water mass transformation, the upgoing limb of the hysteresis loop in Figure 11B (increasing T_{NADW}) is associated with a deeper stratification and a more vigorous lower diffusive cell in the basin, while the upper adiabatic cell is weaker. Steady state conditions in the downgoing limb of the hysteresis loop (decreasing T_{NADW}) is associated with a shallower stratification and a weaker lower diffusive cell in the basin, and a stronger upper adiabatic cell. To transition into the shoaled configuration, interface 2 has to move into the region where κ decreases, causing the diffusive upwelling across the second interface to drop. By deepening this transition region, layer 2 still interacts with the nonlinear portion of the κ profile, when it thickens due to an increase in ϕ. Deepening the transition zone only helps expand the hysteresis window to a point, though, and eventually, z_2^A no longer interacts with the nonlinearity in $\kappa(z)$ and the system behaves as if κ were constant (Figures 9 and 10).

4.3. Circulation Configuration

The key characteristic of a figure-eight configuration is that the deep water formed in the Atlantic is preferentially transferred into the Indo-Pacific where it diffusively upwells and is transferred back to the Atlantic in the upper layers to close the overturning circulation. Therefore, to assess which combinations of T_{NADW} and ϕ lead to two-cell or figure-eight circulation configurations, consider the sum of T_{zonal} in layers 3 and 4. If this quantity is large, then deep water from the Atlantic is transported to the Pacific to upwell, and circulation is more figure-eight-like. If this quantity is small, then deep water from the Atlantic upwells in the

Atlantic, and circulation is more two-cell-like (Figure 11a). By this metric circulation configuration in the model exists as a continuum between figure-eight and two-cell.

Circulation configuration depends both on T_{NADW} and ϕ. A strong formation rate of dense NADW is the most figure-eight-like, and a weaker formation rate of light NADW is the most two-cell-like (Figure 11a). However, as ϕ decreases (i.e., NADW becomes less dense), the dependence of T_{zonal} in layers 3 and 4 on NADW strength decreases. In the real ocean, ϕ can be thought of as the strength of the density contrast between NADW and AABW. A ϕ value of 1 means that NADW and AABW are more similar in density, as in the modern ocean. During the LGM, evidence from various paleoceanographic tracers indicates that AABW was much more dense than NADW (resulting in a shoaled version of NADW), and there was greater stratification at middepth (Adkins, 2002; Curry & Oppo, 2005; Lund et al., 2011), akin to a ϕ value of 0.

According to our metric, the modern ocean, with its dense and strong NADW (~19 Sv, $\phi = 1$), has a a figure-eight circulation configuration. At the LGM, circulation may have been nearly as strong as the modern, but NADW was less dense ($\phi = 0$). Our metric predicts that this would lead to a two-cell circulation configuration, matching paleo observations (Burke et al., 2015; Curry & Oppo, 2005; Skinner et al., 2017). During MIS 3, there is evidence that NADW strength fluctuated and there are abrupt jumps in NGRIP δ^{18}O, indicating rapid flips between states in the climate system (Figure 2). The abrupt jumps in outcrop position and hysteresis that we see for many values of ϕ over a range of NADW fluxes is reminiscent of this period. It is somewhat surprising, however, that the large stratification changes that occur in the hysteresis region (Figure 11b) do not lead to large changes in circulation configuration (Figure 11a). Our configuration metric predicts that MIS 3 had an intermediate circulation structure between the two-cell and figure-eight end members. Other studies have also found that intermediate climate states are more conducive to abrupt climate transitions (Brook & Buizert, 2018; Buizert & Schmittner, 2015; Dome Fuji Ice Core Project Members: et al., 2017; McManus et al., 1999).

While the hysteresis does not seem to have a strong effect on the circulation configuration (Figure 11a), or at least the zonal transport between basins, it does impact the basin stratification, parameterized here as the ratio of the thickness of the upper two layers divided by the thickness of the lower two layers—a shallower stratification has a small ratio and a deeper stratification has a large ratio (Figure 11b). This stratification metric potentially corresponds better with the paleoceanographic observations of Oppo and Curry (2012), Curry and Oppo (2005), and Lund et al. (2011), who note changes in the structure of ocean properties associated with rearrangements of water masses. As can be seen in Figure 11, circulation configuration (i.e., "figure-eight-ness") and ocean stratification are related but do not exactly match each other.

4.4. Role of Sea Ice in Circulation Configuration Changes

Ferrari et al. (2014) highlights the connection between the quasi-permanent sea ice position and the splitting of upper and lower meridional overturning circulation cells in the modern ocean, and uses this relationship to suggest how the ocean may have changed in the past. Expanded sea ice and shoaled NADW relative to AABW is suggested to lead to an overturning configuration with enhanced stratification at middepth and greater separation between the upper and lower circulation cells as the boundary between the cells moves away from the region with rough bottom topography and therefore enhance diffusive mixing. Our model results can be compared with two aspects of Ferrari et al. (2014): (i) the role of rough topography in determining the circulation configuration and (ii) the role of the meridional sea ice extent in determining the circulation configuration.

Ferrari et al. (2014) note that vertical mixing is enhanced near topography and thus intensified below 2,000 m in the ocean (Polzin et al., 1997). To address this feature, our model has a kinked $\kappa(z)$ profile with enhanced vertical diffusivity in the deep ocean. This nonlinear vertical diffusivity distribution is crucial for generating abrupt transitions in basin stratification and Southern Ocean isopycnal outcrop position in this model. Abrupt changes in circulation structure, defined by more of less zonal exchange between Atlantic and Indo-Pacific basins, are less clear (Figure 11a).

Ferrari et al. (2014) also argue that expanded sea ice during the LGM led to a two-cell circulation configuration with shoaled NADW and enhanced stratification between the upper and lower circulation cells. There is an important dynamical step between sea ice expansion and the achievement of two-cell circulation configuration—the authors note that if sea ice is expanded, any NADW that sinks below the 2,000 m becomes part of the lower cell, and this deep penetration of NADW in a glacial configuration would only

be possible transiently (since it would effectively drain the upper cell). They assert that the system would eventually adjust such that NADW was shoaled relative to AABW, since any NADW that was dense enough to sink below 2,000 m would become part of the lower cell leaving the residual NADW less dense. In our model, NADW density is a separate parameter, so we cannot test the full Ferrari et al. (2014) mechanism exactly as they pose it in the paper (since ϕ does not evolve during a simulation). We can, however, test the two parts of this mechanism separately: the effect of expanded sea ice given a deep NADW of constant density ($\phi = 1$), and the effect of shoaled NADW given constant sea ice extent. We find that the density of NADW is more important for determining circulation configuration than the meridional extent of sea ice alone (see Figure S5).

4.5. Limitations of the Model

In addition to the simplifications and idealizations that are discussed in section 2 and the Appendix, there are some additional caveats to the model that we describe here. Despite the insights that can be gained from the model's simplicity, there are important limitations that need to be addressed by future work. Perhaps most relevant to the results shown here is the lack of temporally evolving sea ice. A key result of the simulations is that isopycnals tend not to outcrop where surface buoyancy fluxes are weak, that is, in the sea ice covered region. When isopycnal outcrop positions adjust, they change the relative surface outcrop area of the four density classes and, assuming a relationship between layer density and temperature, the mean temperature of the Southern Ocean. These mean Southern Ocean temperature changes would likely also change the extent of Southern Ocean sea ice, but sea ice in our current model cannot respond. Coupling a dynamic sea ice model to the reduced overturning model presented here is the focus of ongoing work. While dynamic sea ice will likely impact the results presented here, we do not believe that isopycnals will move exactly in sync with dynamic sea ice—changes in NADW formation require changes in water mass transformation at the surface of the Southern Ocean, which in turn require changes in isopycnal outcrop position between regions of positive and negative surface buoyancy flux.

Another caveat of the model is that layers have fixed densities that cannot evolve in time and that are not explicitly broken up into temperature and salinity components. The interplay of temperature and salinity in the equation of state of seawater can lead to interesting dynamics, such as abrupt thermobaric convection (Adkins et al., 2005; Su et al., 2016a, 2016b), and has been linked to multiple steady overturning circulation states (Stommel, 1961). It is also important to include the atmosphere in order to fully capture global dynamics and feedbacks across abrupt climate change events. The atmosphere responds much more quickly than the ocean (Liu & Alexander, 2007) and plays an important role in interhemispheric communication of climate signals. A dynamic atmosphere would also influence Southern Ocean sea ice extent and ocean density via input of heat and precipitation. In addition to these physical effects, atmospheric CO_2 also strongly impacts global temperature and the biogeochemistry of the ocean-atmosphere system.

4.6. Implications for Paleoceanography

Despite the simplifications of our model, we can use the results presented here to speculate about how changes in NADW density and Southern Ocean sea ice extent could have led to the sequence of DO events we observe from the last glacial cycle. Throughout this time, the frequency and character of DO events changed dramatically—they were rare during the end of MIS 5 into the beginning of MIS 4, widespread during MIS 3 and absent during MIS 2 and the interglacials (Figure 2). The frequency and duration of these DO events can be related to the width of the hysteresis loop—if the loop is wide, then there is a larger region with multiple steady states and the climate system can rapidly switch between these states as it gradually drifts to a place where only one steady state is stable. If the hysteresis loop is narrow, then it is easier for the system to drift away from the hysteresis window, making the system more stable.

The first DO events emerge during the end of MIS 5. These events are long in duration but infrequent, characteristic of a regime with a narrow hysteresis loop (Figure 2). Our experiments indicate that the conditions most conducive to narrow hysteresis loops are when sea ice is contracted to the south and NADW is dense (location 1 in Figure 6). These long infrequent DO events persist until the middle of MIS 4, when they nearly disappear. During the middle and end of MIS 4 atmospheric pCO_2 also drops by almost 50 ppm (Figure 2a). This could have caused an expansion of sea ice in the Southern Ocean, thus pushing the hysteresis window to lower NADW flux values (location 2 in Figure 6) and causing the near disappearance of DO events—note that location 2 in panel (b) is just outside the hysteresis window. Moving from MIS 4 to MIS 3, there is a reemergence of DO events, and through MIS 3, these events become more frequent and shorter in duration

(Figure 2), indicating that the hysteresis is widening during this time interval. This behavior is likely caused primarily by the shoaling of NADW in the water column, as hysteresis width is more sensitive to NADW density than sea ice extent, though Southern Ocean sea ice may have been simultaneously expanding (location 3 in Figure 6). Finally, moving into MIS 2 and the Last Glacial Maximum, DO events disappear. This is consistent with the further shoaling of NADW in the water column, which eventually prevents hysteresis (location 4 in Figure 6). By the LGM there is additional evidence from paleo records that NADW is shoaled relative to AABW and that circulation is two-cell rather than figure-eight (Curry & Oppo, 2005; Ferrari et al., 2014; Lund et al., 2011; Roberts et al., 2016; Sikes et al., 2017). These paleo observations are also supported by our circulation configuration metric, which indicates that a fully shoaled NADW would lead to a two-cell circulation configuration (Figure 11).

5. Conclusions

We have presented a two-basin time-dependent dynamical box model of the ocean, incorporating realistic, dynamical characteristics of water mass transformation in the Southern Ocean. We find that across a range of external parameters, small perturbations in the strength of NADW formation can lead to abrupt transitions in the strength and structure of the global overturning circulation. These changes also exhibit that hysteresis as a much larger change in NADW formation is required to return the overturning circulation to its original state. The physical mechanism for the abrupt transitions depends on the unique properties of the Southern Ocean: it can be a site of both positive and negative water mass transformation. The distribution of the surface buoyancy flux is strongly constrained by the cycle of sea ice formation and melt, which results in spatially separated negative and positive buoyancy fluxes. Despite the centrality of the Southern Ocean in the physics driving hysteresis in our model, we find that the distribution of vertical diffusivity in the ocean basins is important as well. Experiments with constant or linear vertical diffusivity profiles do not show hysteresis in this particular model configuration.

We find that the relative density of NADW is the most important factor for determining overturning circulation configuration (two-cell versus figure-eight). Dense NADW and strong NADW formation favors figure-eight circulation while light NADW favors two-cell circulation. The combination of NADW strength and density that gives rise to multiple steady states and abrupt transitions in Southern Ocean outcrop position and basin stratification, characteristic of MIS 3, has an intermediate circulation configuration between pure figure-eight and two-cell. While the two limbs of a hysteresis loop have similar circulation configurations, they involve very different basin stratifications, defined as the ratio of the average thickness of upper ocean layers divided by the average thickness of deep ocean layers.

The duration of abrupt climate transitions, such as the DO events of MIS 3, are influenced by hysteresis width. In our model, we find that the width of the hysteresis loop is dependent on NADW density and sea ice extent. These results have important implications for paleoceanography, as they specify which ocean conditions would favor frequent short DO events or infrequent long DO events and allow us to make inferences about the patterns of sea ice and NADW density change across the last glacial. In the future, it would be beneficial to validate the results of our experiments and the mechanisms we propose in more complex models that represent a greater range of physical processes.

Appendix A: Detailed Model Description

A1. Diffusive Upwelling

Following Nikurashin and Vallis (2011), vertical stratification in the basin is in advective-diffusive balance according to

$$wb_z = \frac{\partial}{\partial z}\left(\kappa(z) b_z\right), \tag{A1}$$

where w is the vertical velocity, b_z is the vertical buoyancy gradient, and κ is the diapycnal diffusivity. This equation can be rearranged as a vertical mass flux by dividing both sides by b_z and multiplying by the horizontal area, A:

$$wA = b_z^{-1}\frac{\partial}{\partial z}\left(\kappa(z) b_z\right) A. \tag{A2}$$

For a low-order model, equation (A2) can be parameterized as

$$T_{\text{diff}}^{iA,P} = \frac{\kappa(z_i^{A,P})A}{\Delta z} \tag{A3}$$

where $T_{\text{diff}}^{iA,P}$ is the vertical diffusive mass flux across layer i in either the Atlantic (A) or the Pacific (P). The parameter Δz is the thickness of the layer above the interface. Models with a larger number of layers could describe this diffusive flux by using more accurate discretization schemes, such as centered difference (e.g., Marshall & Zanna, 2014).

The diapycnal diffusivity varies with depth and is evaluated at the interface depth, $z_i^{A,P}$, according to the equation:

$$\kappa(z_i^{A,P}) = \kappa_0 - \frac{\Delta\kappa}{2}\tanh\left(\frac{z_i^{A,P} - z_0}{d}\right), \tag{A4}$$

where κ_0 is the mean diapycnal diffusivity, $\Delta\kappa$ is the change in diapycnal diffusivity over depth interval d, and z_0 is the depth of the transition. Diapycnal diffusivity is enhanced in the deep ocean due to interaction with rough bottom topography (Polzin et al., 1997; Waterhouse et al., 2014), which has been implicated as an important mechanism for changes in ocean circulation configuration during glacial times (Ferrari et al., 2014).

A2. Zonal Transport

Zonal flow in the Southern Ocean has a mean (barotropic) component and a baroclinic component which is related to the relative slopes of the isopycnals in the ACC. While the zonal transport of the ACC is roughly 150 Sv in the modern ocean, there is a smaller component, O(10 Sv), that represents a convergent zonal transport between the Atlantic and Indo-Pacific basins that is needed to close the overturning. Thompson et al. (2016) accounted both barotropic and baroclinic components of the ACC to calculate the zonal convergence for each layer (e.g., between a layer in the Atlantic and the corresponding layer in the Pacific). Here, we simplify this expression by assuming a barotropic zonal flow, $U_0 = 0.05$ m/s, and convergence or divergence of mass is controlled only by the difference in layer thickness of a given density class between the Atlantic and Indo-Pacific basins. Therefore we define zonal transport as

$$T_{\text{zonal}}^{i} = -U_0\left(A_{\text{ACC}}^{\text{iA}} - A_{\text{ACC}}^{\text{iP}}\right), \tag{A5}$$

where $A_{\text{ACC}}^{\text{iA,P}}$ is the cross-sectional area (in the depth-latitude plane) of a layer in the ACC. The zonal transport is dependent on the difference in cross-sectional area between the Atlantic and Pacific basins, $\Delta A_{\text{ACC}}^{i} \equiv A_{\text{ACC}}^{\text{iA}} - A_{\text{ACC}}^{\text{iP}}$. Total zonal transport for a given layer, ΔT_{zonal} is the sum of the zonal transport in that layer and the zonal transports for all layers above. Zonal transport is defined as positive from Pacific to Atlantic.

A3. Southern Ocean Buoyancy Fluxes and Residual Transport

Buoyancy fluxes in the Southern Ocean are prescribed to represent both sea ice and one due to atmospheric heating. Sea ice formation and brine rejection near the Antarctic continent leads to a strong negative buoyancy flux; freshwater from sea ice melt leads to a positive buoyancy flux. The importance of sea ice dynamics on the meridional overturning circulation has recently been diagnosed from both the Southern Ocean State Estimate (Abernathey et al., 2016) and from observations (Pellichero et al., 2018). Abernathey et al. (2016) show that negative buoyancy fluxes from sea ice formation are concentrated around the Antarctic continent, particularly in the Weddell and Ross Seas (key known sites of bottom water formation). Freshwater fluxes from sea ice melt are located further to the north in a band at roughly 60°S (Abernathey et al., 2016). North of the sea ice edge, direct heating from the atmosphere provides an additional heat flux, but its contribution to the overall buoyancy flux budget for the Southern Ocean is smaller than that of sea ice (Abernathey et al., 2016; Sun et al., 2016). The buoyancy fluxes profile is defined as

$$F_{\text{atm}}^{iA,P} = \frac{h_m}{\tau_r}\left(\hat{b} - b^*\right), y_{\text{si}} < y_i^{A,P} < 0, \tag{A6}$$

$$\hat{b} = \Delta b_{\text{atm}}\left(\frac{y_i^{A,P} - y_{\text{si}}}{\ell}\right), \tag{A7}$$

$$F_{\mathrm{SI}}^{iA,P} = \gamma_N F_0 \exp\left[-\left(\frac{y_i^{A,P} - y_{\mathrm{si}}}{a_N}\right)^2\right] - r_b F_0 \exp\left[-\left(\frac{y_i^{A,P} + \ell}{a_S}\right)^2\right], -\ell < y_i^{A,P} < 0, \tag{A8}$$

$$F_b\left(y_i^{A,P}\right) = F_{\mathrm{atm}}^{iA,P} + F_{\mathrm{SI}}^{iA,P}, \tag{A9}$$

where y_{si} is the meridional position of the sea ice edge in the Southern Ocean, $\hat{b}$ is the meridional buoyancy distribution for the atmospheric component of the buoyancy forcing, Δb_{atm} is the range of atmospheric buoyancy forcing (chosen such that the northern boundary of the ACC corresponds to an atmospheric heat flux of 50 W/m^2), $b^* = (b_n + b_{n+1})/2$ is the buoyancy associated with the interface, h_m is the depth of the mixed layer in the Southern Ocean, and τ_r is the relaxation time scale (set to be equal to 2 months).

In equation (A8), the prefactor γ_N sets the strength of buoyancy forcing in the sea ice melting and formation regions (set to a value of 0.4), and a_N and a_S are the widths of the sea ice melting and formation regions, respectively. The prefactor r_b sets the relative strength of bottom water formation in the Atlantic and Pacific basins; the value is r_b in the Atlantic, and $\left(1 + L_A/L_P\left(1 - r_b\right)\right)$ in the Pacific. This accounts for the difference in Atlantic and Pacific basin widths and ensures that the zonally integrated negative buoyancy flux is equal to the zonally-integrated positive buoyancy flux. This expression also sets limits for r_b of $0 < r_b < 13/3$.

According to the residual-mean theory of Marshall and Radko (2003), the residual transport at the base of the mixed layer in the ACC is defined as $B_0/\partial_y b_0$—the net surface buoyancy flux over the meridional surface buoyancy gradient. In our model we use the net surface buoyancy flux, F_b, as defined in equations (A6)–(A9), and follow the simplification of Thompson et al. (2016), approximating the meridional buoyancy gradient as: $\Delta b/\ell^{ml}$ where $\Delta b = b_n - b_{n+1}$ and $\ell^{ml} = (y_{n-1} - y_{n+1})/2$. Buoyancy-driven transport in the Southern Ocean, T_{res}, is therefore defined as

$$T_{res}^{iA,P} = F_b^{iA,P}\left(\frac{\ell^{ml}}{\Delta b}\right) L_{A,P}. \tag{A10}$$

In a continuous model (as in the real ocean), the residual transport in the Southern Ocean is driven by the surface buoyancy forcing and surface buoyancy gradient at each given point. In our reduced order model, we had to make a choice as to whether we would approximate the residual transport for a discrete number of layers in the same fashion as a continuous model (i.e., at the interface) or by integrating the surface buoyancy forcing over the outcrop area. The integration approach works for layers 2 and 3 but fails for layers 1 and 4—the integrated buoyancy flux for the heaviest and lightest layers is negative and positive, respectively, which would imply residual transport into heavier and lighter layers, which do not exist. Therefore, we calculate $T_{\mathrm{res}}^{iA,P}$ based on $F_b^{A,P}$ evaluated at the interface $y_i^{A,P}$. Because we are tracking the volume budget in the model, and thus water mass transformations between density classes, it makes sense to evaluate these transformations at the interfaces between density classes.

A4. Evolution Equations for y and z

In the Atlantic and Pacific basins, isopycnal depths are defined as a balance between diffusive upwelling and the combination of buoyancy-driven exchange and zonal convergence/divergence in the Southern Ocean. In the Atlantic basin, NADW formation also influences isopycnal depths. Interface depths in the basin evolve according to the general equation:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{1}{A}\left[\Delta T_{\mathrm{diff}} + \Delta T_{\mathrm{res}} + \Delta T_{\mathrm{zonal}} + \left(T_{\mathrm{NADW}}\right)\right], \tag{A11}$$

where A is the horizontal area of the basin and the ΔTs refer to the net volume transport for a given layer. In our model, T_{NADW} is an external parameter and provides the main external forcing. The full NADW flux, T_{NADW}, is added to the evolution equation for z_1, and multiplied by a prefactor ϕ in the evolution equation for z_2. This prefactor determines the relative amount of the total NADW flux that is added to layer 2 versus layer 3. A ϕ value of 1 means that all of T_{NADW} goes into layer 3 and a ϕ value of 0 means that all of T_{NADW} goes into layer 2.

At the surface in the Southern Ocean, we calculate the isopycnal outcrop position by solving the full residual-mean equation. At steady state, transport in the Southern Ocean, ψ_{res}, is the balance between mean wind-driven flow, $\bar{\psi}$, and eddy-driven flow, ψ^* (Marshall & Radko, 2003). If $\bar{\psi}$ and ψ^* are exactly equal, then there is no overturning transport. However, if the system is not in steady state, this need not hold. Therefore, all three components of the residual circulation ($\bar{\psi}$, ψ^*, and ψ_{res}) can contribute to the isopycnal outcropping position (although ψ_{res} is the only one of the three that actually involves water mass transformation). Outcropping positions evolve according to the following equation:

$$\frac{dy}{dt} = \frac{1}{h_m}\left(\bar{\psi} + \psi^* - \psi_{res}\right) = \frac{1}{h_m}\left(-\frac{\tau}{\rho_0 f} + Ks - \frac{T_{res}}{L_{A,P}}\right), \tag{A12}$$

where h_m is the depth of the mixed layer, τ is the wind stress (held at a constant value of τ_0 for the experiments presented here), ρ_0 is the reference density, f is the coriolis parameter, K is the along-isopycnal diffusivity, s is the isopycnal slope, and T_{res} is as defined in equation (A10).

The evolution equations for y and z are solved using a fourth-order Runge-Kutta method. For an experiment where the model is run out to steady state, the model is first initialized for 10 years using a time step of 0.1 days, then run with a time step of 0.5 days for the rest of the simulation (10,000 years).

Acknowledgments
We would like to thank Raffaele Ferrari, Emily Newsom, Andrew Stewart, David Marshall, James Rae, and Andrea Burke for helpful discussions, and two anonymous reviewers, whose comments improved the manuscript. S. K. V. H. received support from NSF grants OCE-1503129 and OCE-1204211 and the Lamont-Doherty Earth Observatory Postdoctoral Fellowship. A. F. T. received support from the David and Lucille Packard Foundation and from NSF grant OCE-1235488. J. F. A. received support from NSF grants OCE-1503129, OCE-1737404, and OCE-1450528. The model code used for this paper will be available on GitHub at the https://github.com/shiness11/DynBoxTwoB4L website.

References

Abernathey, R. P., Cerovecki, I., Holland, P. R., Newsom, E., Mazloff, M., & Talley, L. D. (2016). Water-mass transformation by sea ice in the upper branch of the Southern Ocean overturning. *Nature Geoscience*, *9*(8), 596–601.

Adkins, J. F. (2002). The salinity, temperature, and $\delta^{18}O$ of the glacial deep ocean. *Science*, *298*(5599), 1769–1773.

Adkins, J. F., Ingersoll, A. P., & Pasquero, C. (2005). Rapid climate change and conditional instability of the glacial deep ocean from the thermobaric effect and geothermal heating. *Quaternary Science Reviews*, *24*, 581–594.

Alley, R. B., & Clark, P. U. (1999). The deglaciation of the Northern Hemisphere: A global perspective. *Annual Review of Earth and Planetary Sciences*, *27*, 149–182.

Andersen, K. K., Svensson, A., Johnsen, S. J., Rasmussen, S. O., Bigler, M., Rothlisberger, R., et al. (2006). The Greenland ice core chronology 2005, 15–42 ka. Part 1: Constructing the time scale. *Quaternary Science Reviews*, *25*(23-24), 3246–3257.

Bassis, J. N., Petersen, S. V., & Mac Cathles, L. (2017). Heinrich events triggered by ocean forcing and modulated by isostatic adjustment. *Nature*, *542*(7641), 332–334.

Bereiter, B., Eggleston, S., Schmitt, J., Nehrbass-Ahles, C., Stocker, T. F., Fischer, H., et al. (2015). Revision of EPICA Dome C CO_2 record from 800 to 600 kyr before present. *Geophysical Research Letters*, *42*, 542–549. https://doi.org/10.1002/2014GL061957

Birchfield, G. E., & Broecker, W. S. (1990). A salt oscillator in the glacial Atlantic? 2. A "scale analysis" model. *Paleoceanography*, *5*(6), 835–843.

Böhm, E., Lippold, J., Gutjahr, M., Frank, M., Blaser, P., Antz, B., et al. (2015). Strong and deep Atlantic meridional overturning circulation during the last glacial cycle. *Nature*, *517*(7532), 73–76.

Bond, G., Broecker, W., Johnsen, S., McManus, J., Labeyrie, L., Jouzel, J., & Bonani, G. (1993). Correlations between climate records from North Atlantic sediments and Greenland ice. *Nature*, *365*, 143–147.

Bond, G. C., Showers, W., Elliot, M., Evans, M., Lotti, R., Hajdas, I., et al. (1999). The North Atlantic's 1–2 Kyr climate rhythm: Relation to Heinrich events, Dansgaard/Oeschger cycles and the little ice age. In *Mechanisms of global climate change at millennial time scales* (pp. 35–58). Washington, DC: American Geophysical Union.

Broecker, W. S., Bond, G. C., Klas, M., Bonani, G., & Wolfli, W. (1990). A salt oscillator in the glacial Atlantic? 1. The concept. *Paleoceanography*, *5*(4), 469–477.

Brook, E. J., & Buizert, C. (2018). Antarctic and global climate history viewed from ice cores. *Nature*, *558*, 200–208.

Buizert, C., & Schmittner, A. (2015). Southern Ocean control of glacial AMOC stability and Dansgaard-Oeschger interstadial duration. *Paleoceanography*, *30*, 1595–1612. https://doi.org/10.1002/2015PA002795

Burke, A., Stewart, A. L., Adkins, J. F., Ferrari, R., Jansen, M. F., & Thompson, A. F. (2015). The glacial mid-depth radiocarbon bulge and its implications for the overturning circulation. *Paleoceanography*, *30*, 1021–1039. https://doi.org/10.1002/2015PA002778

Cane, M., & Clement, A. C. (1999). *A role for the tropical Pacific coupled ocean-atmosphere system on Milankovitch and millennial timescales. Part II: Global impacts*, Geophysical Monograph Series (Vol. 112). Washington, DC: American Geophysical Union.

Clement, A. C., & Cane, M. (1999). *A role for the tropical Pacific coupled ocean-atmosphere system on Milankovitch and millennial timescales. Part I: A modeling study of tropical Pacific variability*, GeophysicalMonograph Series (Vol. 112). Washington, DC: American Geophysical Union.

Curry, W. B., & Oppo, D. W. (2005). Glacial water mass geometry and the distribution of $\delta^{13}C$ of ΣCO_2 in the western Atlantic Ocean. *Paleoceanography*, *20*, PA1017. https://doi.org/10.1029/2004PA001021

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., et al. (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, *364*(6434), 218–220.

de Lavergne, C., Madec, G., Roquet, F., Holmes, R. M., & McDougall, T. J. (2017). Abyssal ocean overturning shaped by seafloor distribution. *Nature*, *551*, 181–186.

Dome Fuji Ice Core Project Members:, Kawamura, K., Abe-Ouchi, A., Motoyama, H., Ageta, Y., Aoki, S., Azuma, N., et al. (2017). State dependence of climatic instability over the past 720,000 years from Antarctic ice cores and climate modeling. *Science Advances*, *3*(2), e1600446.

EPICA Community Members (2006). One-to-one coupling of glacial climate variability in Greenland and Antarctica. *Nature*, *444*, 195–198.
Ferrari, R., Jansen, M. F., Adkins, J. F., Burke, A., Stewart, A. L., & Thompson, A. F. (2014). Antarctic sea ice control on ocean circulation in present and glacial climates. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), 8753–8758.
Ferreira, D., Marshall, J., & Rose, B. (2011). Climate determinism revisited: Multiple equilibria in a complex climate model. *Journal of Climate*, *24*, 992–1012.
Gnanadesikan, A. (1999). A simple predictive model for the structure of the oceanic pycnocline. *Science*, *283*(5410), 2077–2079.
Henry, L. G., McManus, J. F., Curry, W. B., Roberts, N. L., Poitrowski, A. M., & Keigwin, L. D. (2016). North Atlantic Ocean circulation and abrupt climate change during the last deglaciation. *Science*, *353*, 470–474.
Jansen, M. F. (2017). Glacial ocean circulation and stratification explained by reduced atmospheric temperature. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(1), 45–50.
Kuhlbrodt, T., Griesel, A., Montoya, M., Hofmann, M., & Rahmstorf, S. (2007). On the driving processes of the Atlantic meridional overturning circulation. *Reviews of Geophysics*, *45*, RG2001. https://doi.org/10.1029/2004RG000166
Lippold, J., Grützner, J., Winter, D., Lahaye, Y., Mangini, A., & Christl, M. (2009). Does sedimentary 231Pa/230Th from the Bermuda Rise monitor past Atlantic meridional overturning circulation? *Geophysical Research Letters*, *36*, 7. https://doi.org/10.1029/2009GL038068
Liu, Z., & Alexander, M. (2007). Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Reviews of Geophysics*, *45*, RG2005. https://doi.org/10.1029/2005RG000172
Lumpkin, R., & Speer, K. (2007). Global ocean meridional overturning. *Journal of Physical Oceanography*, *37*(10), 2550–2562.
Lund, D. C., Adkins, J. F., & Ferrari, R. (2011). Abyssal Atlantic circulation during the last glacial maximum: Constraining the ratio between transport and vertical mixing. *Paleoceanography*, *26*, PA1213. https://doi.org/10.1029/2010PA001938
Lynch-Stieglitz, J. (2017). The Atlantic meridional overturning circulation and abrupt climate change. *Annual Review of Marine Science*, *9*, 83–104.
MacAyeal, D. R. (1993). Binge/purge oscillations of the laurentide ice sheet as a cause of the North Atlantic's Heinrich events. *Paleoceanography*, *8*(6), 775–784.
Manabe, S., & Stouffer, R. J. (1988). Two stable equilibria of a coupled ocean-atmosphere model. *Journal of Climate*, *1*, 841–866.
Marshall, J., & Radko, T. (2003). Residual-mean solutions for the Antarctic circumpolar current and its associated overturning circulation. *Journal of Physical Oceanography*, *33*(11), 2341–2354.
Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, *5*(3), 171–180.
Marshall, D. P., & Zanna, L. (2014). A conceptual model of ocean heat uptake under climate change. *Journal of Climate*, *27*(22), 8444–8465.
McManus, J. F., Francois, R., Gherardi, J. M., Keigwin, L. D., & Brown-Leger, S. (2004). Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes. *Nature*, *428*(6985), 834–837.
McManus, J. F., Oppo, D. W., & Cullen, J. L. (1999). A 0.5-million-year record of millennial-scale climate variability in the north atlantic. *Science*, *283*, 971–975.
NGRIP members (2004). High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature*, *431*, 147–15.
Nikurashin, M., & Vallis, G. (2011). A theory of deep stratification and overturning circulation in the ocean. *Journal of Physical Oceanography*, *41*(3), 485–502.
Oppo, D. W., & Curry, W. B. (2012). Deep Atlantic circulation during the Last Glacial Maximum and deglaciation. *Nature Education Knowledge*, *3*(10), 1.
Pellichero, V., Sallée, J.-B., Chapman, C. C., & Downes, S. M. (2018). The Southern Ocean meridional overturning in the sea-ice sector is driven by freshwater fluxes. *Nature Communications*, *9*, 1789.
Polzin, K. L., Toole, J. M., Ledwell, J. R., & Schmitt, R. W. (1997). Spatial variability of turbulent mixing in the abyssal ocean. *Science*, *276*, 93–96.
Rahmstorf, S. (2002). Ocean circulation and climate during the past 120,000 years. *Nature*, *419*, 207–214.
Rahmstorf, S., Crucifix, M., Ganopolski, A., Goosse, H., Kamenkovich, I., Knutti, R., et al. (2005). Thermohaline circulation hysteresis: A model intercomparison. *Geophysical Research Letters*, *32*, L23605. https://doi.org/10.1029/2005GL023655
Rasmussen, S. O., Andersen, K. K., Svensson, A. M., Steffensen, J. P., Vinther, B., Clausen, H. B., et al. (2006). A new Greenland ice core chronology for the last glacial termination. *Journal of Geophysical Research*, *111*, D06102. https://doi.org/10.1029/2005JD006079
Roberts, J., Gottschalk, J., Skinner, L. C., Peck, V. L., Kender, S., Elderfield, H., et al. (2016). Evolution of South Atlantic density and chemical stratification across the last deglaciation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 514–519.
Sarnthein, M., Winn, K., Jung, S. J. A., Duplessy, J.-C., Labeyrie, L., Erlenkeuser, H., & Ganssen, G. (1994). Changes in east Atlantic deepwater circulation over the last 30,000 years: Eight time slice reconstructions. *Paleoceanography*, *9*(2), 209–267.
Shakespeare, C. J., & Hogg, A. M. (2012). An analytical model of the response of the meridional overturning circulation to changes in wind and buoyancy forcing. *Journal of Physical Oceanography*, *42*, 1270–1287.
Sikes, E. L., Allen, K. A., & Lund, D. C. (2017). Enhanced δ13C and δ18O differences between the south Atlantic and south Pacific during the last glaciation: The deep gateway hypothesis. *Paleoceanography*, *32*, 1000–1017. https://doi.org/10.1002/2017PA003118
Skinner, L., Primeau, F., Freeman, E., de la Fuente, M., Goodwin, P. A., Gottschalk, J., et al. (2017). Radiocarbon constraints on the glacial ocean circulation and its impact on atmospheric CO_2. *Nature Communications*, *8*, 16010.
Speer, K., Rintoul, S. R., & Sloyan, B. (2000). The diabatic deacon cell. *Journal of Physical Oceanography*, *30*, 3212–3222.
Stommel, H. (1961). Thermohaline convection with two stable regimes of flow. *Tellus*, *13*(2), 224–230.
Su, Z., Ingersoll, A. P., Stewart, A. L., & Thompson, A. F. (2016a). Ocean convective available potential energy. Part I: Concept and calculation. *Journal of Physical Oceanography*, *46*, 1081–1096.
Su, Z., Ingersoll, A. P., Stewart, A. L., & Thompson, A. F. (2016b). Ocean convective available potential energy. Part II: Energetics of thermobaric convection and thermobaric cabbeling. *Journal of Physical Oceanography*, *46*, 1097–1115.
Sun, S., Eisenman, I., & Stewart, A. L. (2016). The influence of Southern Ocean surface buoyancy forcing on glacial-interglacial changes in the global deep ocean stratification. *Geophysical Research Letters*, *43*, 8124–8132. https://doi.org/10.1002/2016GL070058
Svensson, A., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Davies, S. M., et al. (2008). A 60,000 year Greenland stratigraphic ice core chronology. *Climate of the Past*, *4*, 47–57.
Talley, L. (2013). Closure of the global overturning circulation through the Indian, Pacific, and Southern Oceans: Schematics and transports. *Oceanography*, *26*(1), 80–97.
Thompson, A. F., Stewart, A. L., & Bischoff, T. (2016). A multibasin residual-mean model for the global overturning circulation. *Journal of Physical Oceanography*, *46*(9), 2583–2604.

Vinther, B. M., Clausen, H. B., Johnsen, S. J., Rasmussen, S. O., Andersen, K. K., Buchardt, S. L., et al. (2006). A synchronized dating of three Greenland ice cores throughout the Holocene. *Journal of Geophysical Research*, *111*, D13102. https://doi.org/10.1029/2005JD006921

Waterhouse, A. F., MacKinnon, J. A., Nash, J. D., Alford, M. H., Kunze, E., Simmons, H. L., et al. (2014). Global patterns of diapycnal mixing from measurements of the turbulent dissipation rate. *Journal of Physical Oceanography*, *44*, 1854–1872.

Wolfe, C. L., & Cessi, P. (2011). The adiabatic pole-to-pole overturning circulation. *Journal of Physical Oceanography*, *41*(9), 1795–1810.

Wolff, E. W., Chappellaz, J., Blunier, T., Rasmussen, S. O., & Svensson, A. (2010). Millennial-scale variability during the last glacial: The ice core record. *Quaternary Science Reviews*, *29*, 2828–2838.

Zhang, X., Knorr, G., Lohmann, G., & Barker, S. (2017). Abrupt North Atlantic circulation changes in response to gradual CO_2 forcing in a glacial climate state. *Nature Geoscience*, *10*, 518–523.

Zhang, X., Lohmann, G., Knorr, G., & Purcell, C. (2014). Abrupt glacial climate shifts controlled by ice sheet changes. *Nature*, *512*(7514), 290–294.